42. Porter, N. A. & Weekes, T. C. *Mon. Not. R. astrr. Soc.* **183**, 205-210 (1978).
43. Akerlof, C. W. *et al.* in *Proc. 21st ICRC, Adelaide*, vol. 2 (ed. Protheroe, R. J.) 135-138 (Adelaide University, 1990).
44. Porter, N. A. & Weekes, T. C. *Astrophys. J.* **212**, 224-226 (1977).
45. Porter, N. A. & Weekes, T. C. *Mon. Not. R. Astr. Soc.* **183**, 205-210 (1978).
46. Porter, N. A. & Weekes, T. C. *Nature* **277**, 199 (1979).
47. Nolan, K., Porter, N. A., Fegan, D. J., Chantell, M. & Weekes, T. C. *Proc. 21st ICRC, Adelaide*, vol. 2 (ed. Protheroe, R. J.) 150 (Adelaide University, 1990).
48. Fegan, D. J., McBreen, B., O'Brien, D. & O'Sullivan, C. *Nature* **271**, 731-732 (1978).
49. Bhat, P. N. *et al. Nature* **284**, 433-434 (1984).
50. Porter, N. A. & Weekes, T. C. *Nature* **267**, 500-501 (1977).
51. O'Mongain, E. *Nature* **242**, 136-137 (1973).
52. Huguenin, G. R. & Moore, E. L. *Astrophys. J.* **187**, L57-L58 (1974).
53. Phinney, S. & Taylor, J. H. *Nature* **277**, 117-118 (1979).
54. Rees, M. J. *Nature* **266**, 333-334 (1977).
55. Jelley, J. V., Baird, G. A. & O'Mongain, E. *Nature* **267**, 499-500 (1977).

ACKNOWLEDGEMENTS. We thank C. J. Goebel for giving his time whenever we needed expert help. We thank N. Porter for discussions. This research was supported in part by the University of Wisconsin Research Committee with funds granted by the Wisconsin Alumni Research Foundation, in part by the US Department of Energy, in part by the Xunta de Galicia (Spain) and in part by the Smithsonian Scholarly Studies Research Fund. J.H.M. is a NAS/NRC research associate.

ARTICLES

Crystal structure of insecticidal δ-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution

Jade Li*, Joe Carroll† & David J. Ellar†

* Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK
† Biochemistry Department, Cambridge University, Tennis Court Road, Cambridge CB2 1QW, UK

The structure of the δ-endotoxin from *Bacillus thuringiensis* subsp. *tenebrionis* that is specifically toxic to Coleoptera insects (beetle toxin) has been determined at 2.5 Å resolution. It comprises three domains which are, from the N- to C-termini, a seven-helix bundle, a three-sheet domain, and a β sandwich. The core of the molecule encompassing all the domain interfaces is built from conserved sequence segments of the active δ-endotoxins. Therefore the structure represents the general fold of this family of insecticidal proteins. The bundle of long, hydrophobic and amphipathic helices is equipped for pore formation in the insect membrane, and regions of the three-sheet domain are probably responsible for receptor binding.

THE δ-endotoxins are a family of insecticidal proteins produced by *Bacillus thuringiensis* (B.t.) during sporulation, having relative molecular masses (M_r) 60,000-70,000 (60K-70K) in the active form and specific toxicities against insects in the orders of Lepidoptera, Diptera and Coleoptera[1,2]. These toxins have been formulated into commercial insecticides for three decades[3], and now insect-resistant plants are engineered by transformation with Lepidoptera-specific toxin genes[4-6]. In the bacterium δ-endotoxins are synthesized as protoxins of M_rs 70K-135K and crystallize as a parasporal inclusion ~1 μ in size, in which form they are ingested by the susceptible insect. The microcrystal dissolves in the alkaline pH of the midgut and the protoxin is cleaved by gut proteases to release the active toxin. δ-Endotoxins activated *in vitro* bind specifically and with high affinity ($k_D \approx$ 0.1-20 nM) to protein receptors on brush-border membrane vesicles derived from the gut epithelium of target insects[7-9] and create leakage channels of 10-20 Å diameter in the cell membrane[10]. *In vivo* such membrane lesions lead to swelling and lysis of the gut epithelium[11] and death of the insect ensues through starvation and septicaemia. Active δ-endotoxins of different specificities show five strongly conserved regions in their amino-acid sequences[1,12]. Exchanging sequence segments in the divergent regions between toxins of different specificities can produce active hybrids showing altered target specificity[13-15]. We have determined the atomic structure of a Coleoptera-specific δ-endotoxin (CryIIIA, beetle toxin) from *B.t.* subsp. *tenebrionis*[16-18] to elucidate the structural basis for target specificity and membrane perforation by this family of proteins.

Structure determination

Parasporal crystals of the beetle toxin contain the full-length 644-residue protoxin[17] as the minor component, and a product of bacterial processing with 57 residues removed from the N-terminus as the major component[19]. The latter (M_r 67K) is similar in sequence to the active form of other δ-endotoxins. After solubilization, papain cleavage converts the mixture to the 67K toxin (see legend to Table 1). This was recrystallized in the original crystal form of the parasporal crystals, space group $C222_1$ and cell dimensions 117.1 by 134.2 by 104.5 Å, containing one molecule per asymmetric unit and 55% solvent by volume[18].

Initial evaluation of derivatives was carried out at 4.5 Å resolution with data collected on the FAST TV diffractometer[20] using CuKα radiation. Complete datasets (Table 1) were then collected to 2.5 Å resolution from native crystals using the imaging plate systems at the EMBL outstation at DESY and from the mercury and platinum derivatives on film at SRS Daresbury. The electron density map (Fig. 1) at 2.5 Å resolution calculated with phases from multiple isomorphous replacement (mean figure of merit, 0.63) was easily interpretable and was improved by solvent flattening[21,22]. A continuous polypeptide chain from residue 61 to residue 644 at the C terminus was traced unambiguously, and most side-chain atoms could be located in the map. The atomic model was built using the graphics program O (ref. 23) and had an initial R-factor of 37% for all data to 2.5 Å. After preliminary refinement using the program X-PLOR (ref. 24), the current model, containing 584 amino acid residues and 40 bound water molecules, has an R-factor of 19.9% and r.m.s. bond length deviation of 0.017 Å.

Description of the structure

Overview. The beetle toxin is a wedge-shaped molecule with a radius of gyration of 58 Å. As shown in Fig. 2*a*, it comprises three domains. Domain I, from the N terminus of the 67K toxin to residue 290, is a seven-helix bundle in which a central helix is completely surrounded by six outer helices tilted at about +20° to it (Fig. 3*b*,*c*). Domain II, from residues 291 to 500, contains three antiparallel β sheets packed around a hydrophobic core with a triangular cross-section (Fig. 4). Domain III, from residues 501 to 644 at the C terminus is a sandwich of two antiparallel β sheets (Fig. 5). Domains I and III make up the

TABLE 1 Data collection and phasing statistics

Data collection

Data	Method of collection	Number of crystals	Resolution (Å)	Number of measurements	Unique reflections (% completeness)	R_{merge}
Native	image plate	8	2.5	121,767	27,727 (100)	0.108
CH_3HgNO_3	film	7	2.5	103,623	27,767 (100)	0.095
$Hg(CH_3COO)_2$	film	5	2.5	60,224	25,919 (94.5)	0.103
cis-$Pt(NH_3)_2Cl_2$	film	7	2.5	86,629	25,924 (94.5)	0.107
K_2OsO_4	FAST	1	4.5	21,143	4,680 (100)	0.077
$HoCl_3$	FAST	1	4.5	20,013	4,701 (100)	0.069

Phasing statistics

Derivative	Anomalous data	Number of sites	R_{deriv}†	R_{Cullis}‡	Phasing power§ (resolution, Å)
CH_3HgNO_3	no	3	0.183	0.715	1.56 (2.5)
$Hg(CH_3COO)_2$	yes	6	0.247	0.609	2.28 (2.5)
cis-$Pt(NH_3)_2Cl_2$	no	5	0.185	0.682	1.54 (2.5)
K_2OsO_4	no	4	0.149	0.757	1.26 (5.5)
$HoCl_3$	no	3	0.095	0.741	1.35 (5.0)

Protein preparation: Solubilized parasporal crystals from *B.t.* subsp *tenebrionis* were incubated at 0.5 mg ml^{-1} protein with 0.125 units per ml of Agarose-linked papain (Boehringer) in 3.3 M NaBr, 0.05 M sodium phosphate, pH 7.0, and 0.1 mg ml^{-1} phenylmethylsulphonylfluoride (PMSF) for 30 min at 20 °C. Digestion was stopped by adding tosyl lysinechloromethylketone (TLCK) to 0.125 mg m^{-1} and Na_2CO_3 to one fifth volume and removing the enzyme-beads. The 67K beetle toxin was then purified by gel filtration on Sephadex G75 equilibrated with 0.1 M $NaHCO_3$, pH 10.5, 0.5 M NaBr. **Crystallization:** Single crystals were obtained by microdialysis at a protein concentration of 2.5 mg ml^{-1} against 0.1 M $NaHCO_3$, pH 9.5, 1.2 M NaBr at 4 °C overnight, then against 0.1 M $NaHCO_3$, pH 9.2, 0.5 M NaBr at 16 °C; 3 mM NaN_3, 0.1 mM PMSF and 0.1 mg ml^{-1} TLCK were present in all buffers. Crystals were transferred by stages to 0.05 M 2-(*N*-morpholino)ethanesulphonic acid (MES), pH 6.5, for derivative preparation and mounted in 0.03% low-melting agarose in this buffer during data collection. **Data collection:** Image plate and film data were processed using MOSFLM (Imperial College, London) and CCP4 programs (Daresbury, UK). FAST (ref. 20) data were collected and processed with MADNES[45], and scaled in 3° batches. **Derivatives:** Crystals were soaked respectively in 0.25 mM CH_3HgNO_3 for 3.5 h, in 1 mM $Hg(CH_3COO)_2$ for 14 h, in freshly prepared 1 mM *cis*-$Pt(NH_3)_2Cl_2$ for 21 h, in saturated K_2OsO_4 for 35 h, and in 2 mM $HoCl_3$ for 3 days. **Phase calculation:** Two heavy-atom sites in each derivative were located from difference Patterson functions, except in the case of $Hg(CH_3COO)_2$ for which 3 sites were located, and the remaining sites were found by cross-phased difference Fouriers. Heavy-atom parameters were refined against centric data and phases calculated for all data using the program PHARE (G. Bricogne). The two low-resolution derivatives were refined against phases calculated from the high-resolution derivatives. Phasing with the three high-resolution derivatives gave an overall figure of merit of 0.61 (25–2.5 Å) and a clearly interpretable map. Including the remaining derivatives slightly improved the connectivity of the map (overall figure of merit 0.63), and four cycles of solvent flattening using a 50% solvent content and a 9 Å radius in mask calculation[21,22] improved the overall definition of densities. The starting model was built using the program *O* (ref. 23) with the Bones option for main-chain tracing and the autobuild and manip options for side chains. Refinement by simulated annealing using the program X-PLOR (ref. 24) reduced the *R*-factor from 0.37 to 0.25 without individual *B*-factors, and to 0.23 with restrained individual *B*-factors. The model was adjusted in the loops 154–156, 429–436, and 483–488, and had 40 solvent molecules added, then refined by X-PLOR again. The current model has an *R*-factor of 19.9%, with r.m.s. bond length deviation of 0.017 Å, r.m.s. bond angle deviation of 3.2°, and average atomic *B*-factor of 18 Å^2.

* $R_{merge} = \sum_i \sum_i |I_i - \langle I \rangle| / \sum_i |\langle I \rangle|$, where I_i are intensity measurements for a reflection, and $\langle I \rangle$ is the mean intensity for this reflection.

† $R_{deriv} = \sum |F_{PH} - F_P| / \sum |F_P|$, where F_{PH} is the structure factor amplitude of the derivative crystal and F_P is that of the native.

‡ $R_{Cullis} = \sum ||F_{PH} \pm F_P| - F_H(calc)| / \sum |F_{PH} - F_P|$, where F_P and F_{PH} are defined as for R_{deriv}, and F_H(calc) is the calculated heavy-atom structure factor amplitude summed over centric data only.

§ Phasing power = $\langle F_H \rangle / E$, the r.m.s. heavy-atom structure factor amplitudes divided by the residual lack of closure error.

bulky end of the molecule. Through their contact one of the two β sheets in domain III is almost entirely buried. To our knowledge (see, for example, ref. 25), the packing of helices in domain I and of sheets in domain II are both novel arrangements.

Domain I. The central helix in this seven-helix bundle is α_5 (Fig. 3*b*,*c*), which is oriented with its C terminus towards the bulky end of the molecule. Viewed from this end, the outer helices are arranged anticlockwise in the order of α_1, α_2, α_3, α_4, α_6 and α_7, with helices α_1 and α_7 adjacent to the β-sheet domains; α_2 is interrupted by a non-helical section and only the leading half, α_{2a}, is packed against α_5. Figure 3*a* shows the alignment of amino-acid sequence on the surfaces of the helices. The helices are long, especially α_3 to α_7, which contain respectively 8, 7, 6, 9 and 7 complete helical turns and hence would be long enough to span the 30-Å thick hydrophobic region of a membrane bilayer. Furthermore, the six outer helices bear a strip of hydrophobic residues (defined by $\Delta G \geq 0$ for transfer from oil to water) down their entire length on the side-facing helix α_5, so they are amphipathic. In keeping with the general observation that secondary structures are close-packed and bury hydrophobic surfaces[26], the helix contact angles in this domain cluster around +20° rather than −50°, giving the bundle a bouquet-like appearance (Fig. 3*b*). Figure 3*c* shows the bundle in cross-section. The interhelical space contains 27 aromatic residues which are packed in the edge-to-face fashion[27]; all polar groups in this region are hydrogen-bonded or in salt bridges.

The concentric arrangement of the seven-helix bundle is distinct from the two-layered type seen in bacteriorhodopsin. There is some resemblance to the pore-forming domain of colicin A[28], in which two hydrophobic helices are shielded from solvent by eight amphiphilic helices, but the colicin helices are generally shorter. Like the colicin helices, the bundle in the beetle toxin may be a soluble form of packaging for the hydrophobic and amphiphilic helices that will form pores in the membrane after a large change in conformation.

Domain II. In Fig. 4*a* and 4*b* the three sheets of this domain are laid side-by-side, as they would be seen from the solvent. There is an apparent structural duplication between the four-stranded antiparallel sheets, sheet 1 and sheet 2. The chain connections, β_4, β_3, β_2, β_5 and β_8, β_7, β_6, β_9, respectively, follow the order of +3, −1, −1, +3, which is typical of the 'Greek-key' topology[29]. From both sheets the inner strands, β_3 and β_2 as well as β_7 and β_6, extend some 20 Å to the apex of the molecule as two-stranded β ribbons; and at the point of departure from the sheets there is a β-bulge in β_3 and in β_7 to twist the plane of the ribbon by nearly 90° relative to the sheet. The connections between the outer strands cross over the ribbons on the solvent side.

The pseudo-symmetry between these sheets is very approximate. Using the least squares option in *O* (ref. 23), the sheet region of the strands β_3 and β_2 can be brought to superimpose on that of β_7 and β_6, with a r.m.s. fit of 0.72 Å for 13 α carbons. But the r.m.s. fit increased to 1.1 Å for 23 α carbons of the

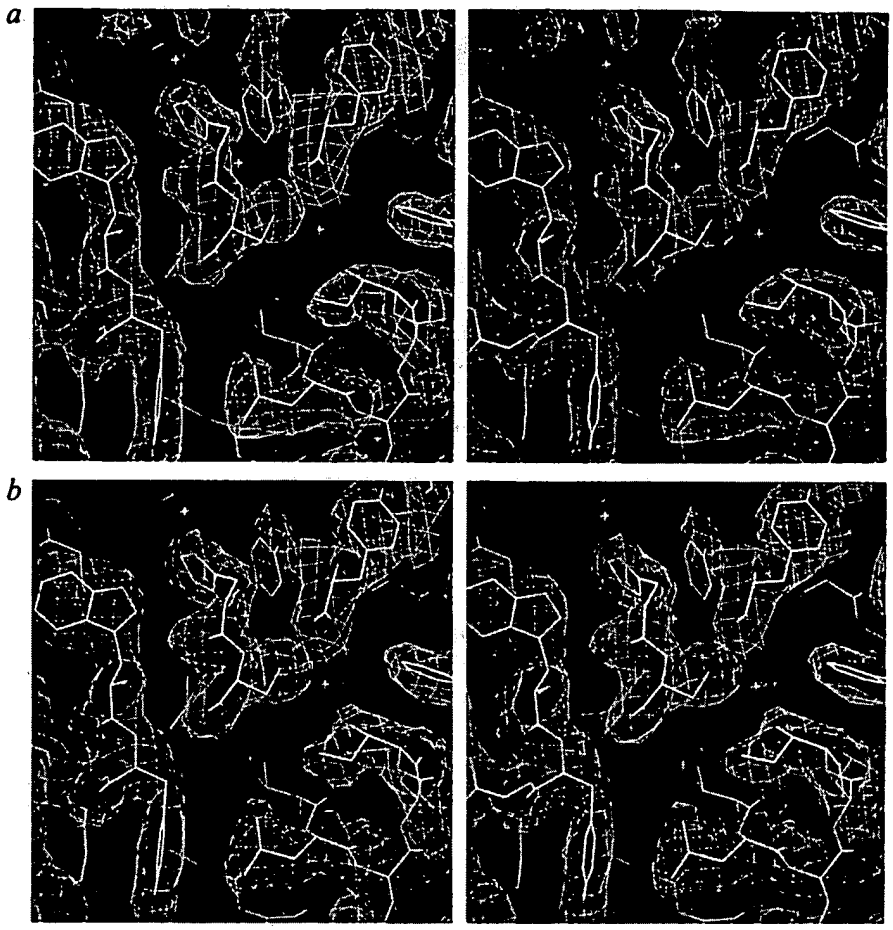


FIG. 1 Electron density map in the neighbourhood of Cys 243, calculated *a*, using combined phases[46] from multiple isomorphous replacement and solvent flattening, and *b*, using combined experimental and model phases[46] after refinement by X-PLOR. The refined structure is shown superimposed for reference. Although Cys 243 is a major site of both the methylmercury (MM) and mercuric acetate (MA) derivatives, the methyl mercury site is in a hydrophobic environment compared with the mercuric acetate site.

whole inner strands including the ribbon region, and 1.7 Å for 36 α carbons on all four strands. Nonetheless, the sequence alignment brought by this superposition of the two sheets revealed a low level of internal homology, with seven pairs of equivalent residues (shown in bold) out of 41 aligned α carbons:

```
338  HRIQPHTRPQP(6)SFNYWS(1)NYVSTRPSI(0)GSNDIITSPF(10)NLKFN  395
402  AVANTNLAVWP(0)SAVYSG(1)TKVEFSQYN(3)DEASTQTYDS(7 )SWDSI  453
```

The three-stranded sheet 3 is formed by two separate polypeptide segments. The C-terminal segment of domain II contributes the two-stranded ribbon of β_{10} and β_{11}, whereas the N-terminal segment of this domain contributes strand β_1, which is hydrogen-bonded to β_{11}; β_1 is followed by a two-turn helix α_8 and an extended chain.

Figure 4*c* and *d* shows in side view and in cross-section that the three antiparallel sheets are packed around a triangular hydrophobic core. This brings the strand β_{10} on the edge of sheet 3 into proximity with strand β_4 on the edge of sheet 1, as well as placing the loops at the end of the three β ribbons into a region of about 12 Å radius at the molecular apex. This domain is in contact with helix α_7 of domain I on the face of sheet 3 (Fig. 4*c*).

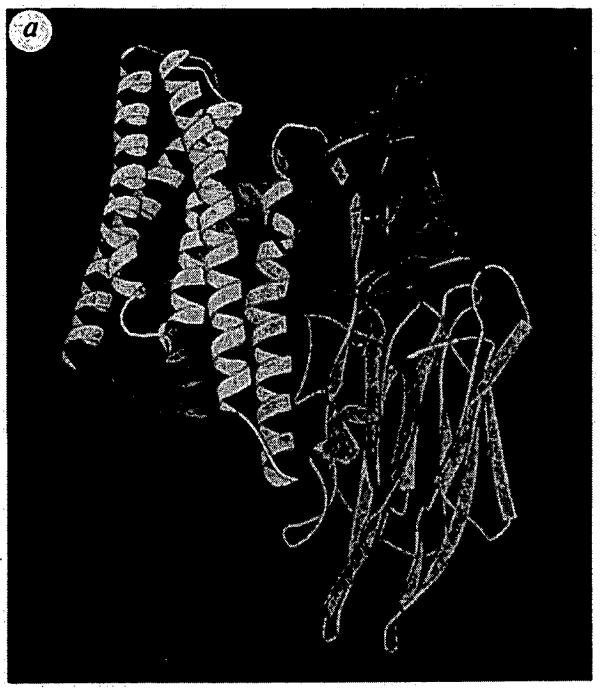
Domain III. Figure 5 is a ribbon drawing of the strands forming the two sheets of the β sandwich. The sheet containing the C-terminal strand is in contact with domain I and will be called the inner sheet. This domain has the 'jelly-roll' topology[29], because it can be generated by folding an antiparallel β ribbon which starts with β_{13} (N terminus) and β_{23} (C terminus) on the inner sheet, and ends in the loop between β_{18} and β_{19} on the outer sheet; β_{14} is a short excursion from this ribbon and forms the fifth antiparallel strand of the outer sheet. In addition, small parallel sheets are formed at the edge of the β sandwich through hydrogen bonding of strand β_{12} to β_{16} at the edge of the outer sheet, and β_1 to β_{13} at the edge of the inner sheet.

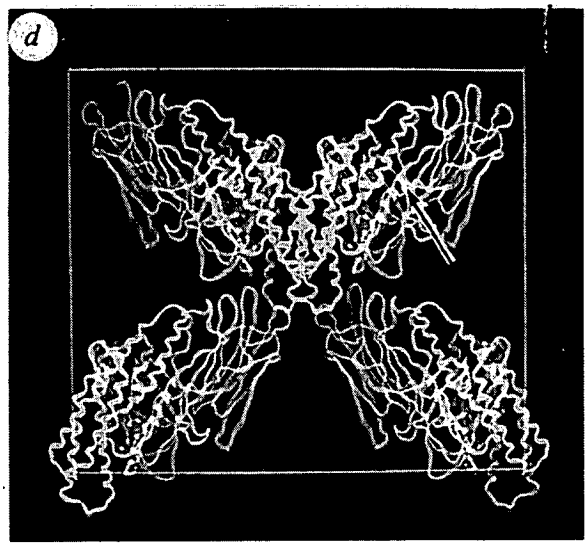
Distribution of conserved sequences. The core of the beetle toxin molecule encompassing the domain interfaces is built from the five sequence blocks that are highly conserved throughout the δ-endotoxin family[1] (Fig. 2*b*,*c*). Block 1, located in the beetle toxin sequence at residues 189–218, corresponds to the central helix (α_5) of the bundle in domain I. Block 2, residues 239–305, overlaps with the latter half of α_6, and with α_7 and β_1; the latter hydrogen-bonds to the edge of the inner sheet in domain III before forming part of the three-stranded sheet 3 in domain II. Block 3, residues 491–538, overlaps with the latter part of β_{11}, where it is hydrogen-bonded to β_1, and with the loops connecting domains II and III. The remainder of block 3 together with blocks 4 and 5, namely residues 560–569 and 633 to the C terminus, respectively, constitute the three buried strands of the inner antiparallel sheet in domain III. The high degree of conservation of internal residues implies that homologous proteins would adopt a similar fold. Using the beetle toxin structure as a model, we can therefore propose a basis for the insecticidal activity of δ-endotoxins as a family.

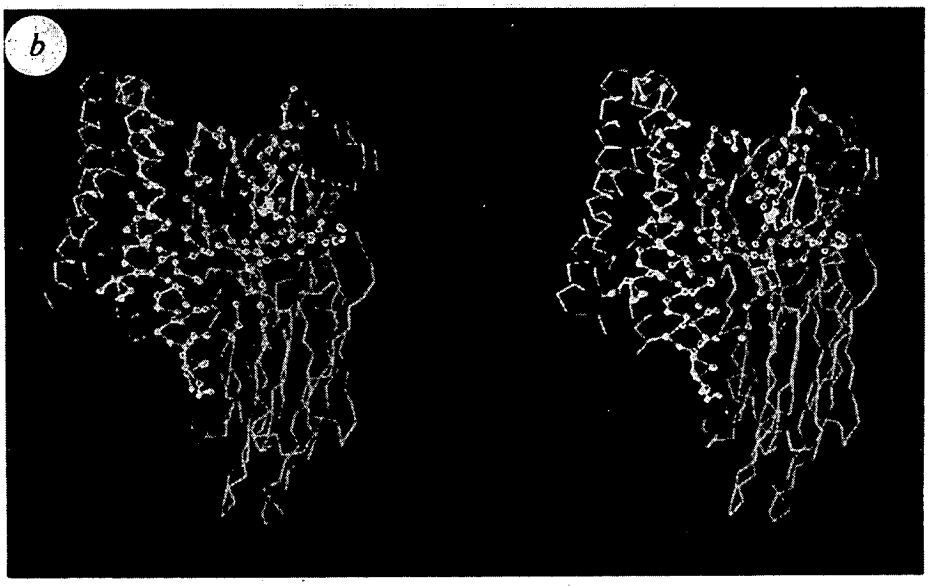
Basis of insecticidal function

Solubility. The beetle toxin crystals are isomorphous with the parasporal crystals[18,19] and show the molecular contacts responsible for solubility behaviour *in vivo*. Four intermolecular salt bridges, Asp 142–Arg 165, Asp 224–Arg 562, Asp 590–Arg 178, and Glu 223–Lys 293, are located at contacts to three different neighbouring molecules. Such salt bridges keep the protoxin crystals insoluble until exposed to the extreme pHs in the insect midgut.

Proteolytic activation. Pro-δ-endotoxins have M_rs of either ~130K or ~70K. Activation by larval gut proteases removes the C-terminal half of the larger protoxins[30,31] and cleaves them at residue 28 or 29 from the N terminus. The smaller protoxins, such as that of the beetle toxin, are processed only at the N







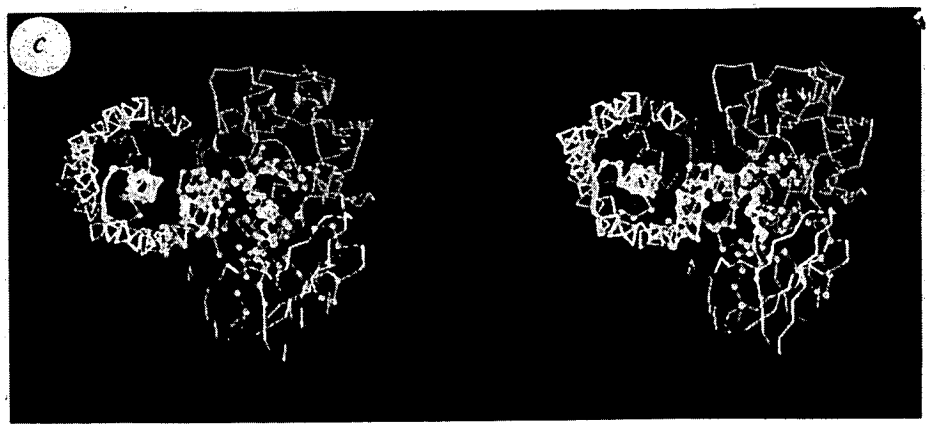


FIG. 2 Overview. *a*, Schematic ribbon representation of the beetle toxin showing the domain organization. Secondary structure assignments are given by Yasspa within program *O* (ref. 23). The polypeptide pathway is indicated by colouring the chain in the rainbow order, from red at the N terminus to blue at the C terminus. The three domains are: I, a seven-helix bundle (upper left); II, a three-sheet assembly (bottom); and III, a β sandwich (upper right). This and all following illustrations of the structure are made with the program MOLSCRIPT[47]. *b* and *c*, Cα trace (stereoview) of the molecule with the five conserved sequence blocks indicated by small beads at their Cα positions. In *b* the view is as in *a*, and in *c* it is down the central helix of the bundle from the bulky end of the molecule; *c* shows that the central helix of domain I and the inner sheet of domain III are conserved; *b* shows that the helices at the domain I–II interface and the loops at the domain II–III interface are also conserved. Note in *c* the helix packing of six around one in domain I. *d*, The solvent channel in the $C222_1$ lattice viewed along the *c* axis. One half of the unit cell thickness is shown, containing four molecules. The other half of the cell is related to this by a two-fold rotation about horizontal axes (blue lines) at $(\frac{1}{2}, y, \pm\frac{1}{4})$. The stacking of both layers leaves solvent channels that traverse the cell along the *c* direction. The N terminus of the molecule (arrow) is accessible from these channels.

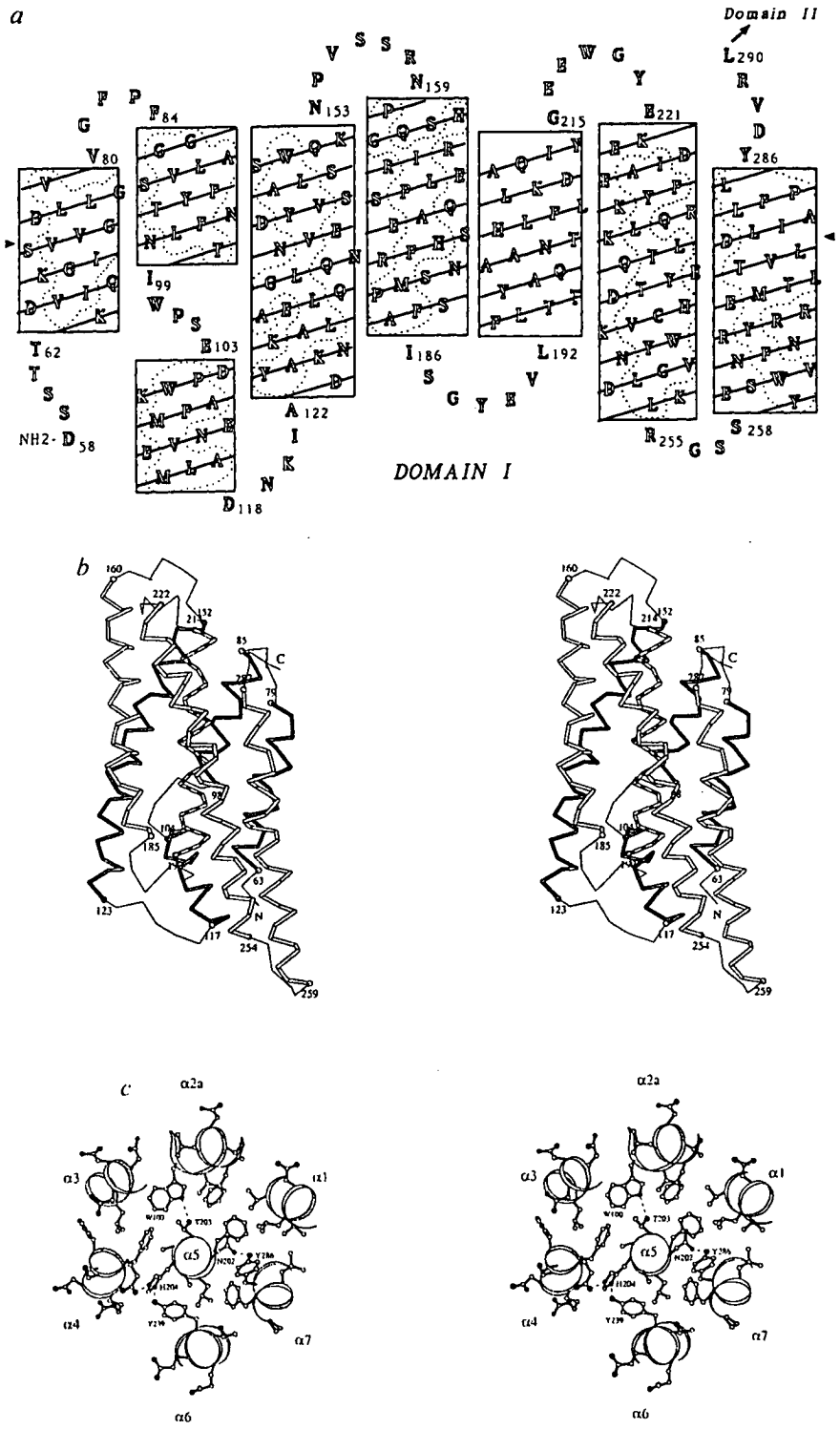


FIG. 3 The seven-helix bundle. *a*, Helical nets showing the position of amino-acid residues along the 7 helices: α_1 (63-79); α_2 (α_{2a}, 85-98 and α_{2b}, 104-117), σ_3 (123-152), α_4 (160-185), α_5 (193-214), α_6 (222-254) and α_7 (259-285). The cylindrical surface of the helices are cut longitudinally on the side facing the solvent and flattened to give a view from the interior of the bundle. The top of the drawing corresponds to the bulky end of the whole molecule. Owing to tilting of the outer helices, different helices are in register vertically only at a level indicated by two arrows pointed at α_1 and α_7; α_5 is the central helix. Dotted curves outline the strip of hydrophobic residues down the inward surface of the other six helices. *b*, Cα trace (stereoview) for the bundle viewed perpendicular to α_5. The relative tilt of the outer helices to α_5 and that between adjacent outer heleices are both about 20°. The Cα trace is shaded grey over helices α1 to α3 in the back, striped over helix α5 in the centre, and white over helices α4, α5, and α7 in the front. *c*, Cross-section of the bundle at the level indicated by the arrows in *a*, viewed from the bulky end of the molecule. The hellical backbone is represented by curly ribbons passing through the Cα positions. The outer helices are positioned roughly hexagonally around the central one and tilted relative to it, so the bundle forms a left-handed superhelix. The aromatic side chains are packed in an edge-to-face fashion. Hydrogen bonds are shown for side-chain atoms.

terminus[19,32] where about 50 residues are removed. The activated δ-endotoxins show a conserved C-terminus, so-called sequence block 5 (ref. 1). Its position as the middle strand of the buried β sheet in domain III precludes further processing from the C terminus. In fact deletion from this site by 4 to 8 residues results in inactive mutants with altered solubility and immunogenicity[30,33-35]. This is not surprising as the inner sheet can be expected to play a critical part in the structural integrity and stability of the toxins through interaction with the helical bundle.

At the N-terminal cleavage sites the different protoxin sequences show locally similar hydropathy profiles[36,37], which would be consistent with a common topology for the N-terminal region of the activated toxins as seen in the helical bundle of the beetle toxin. In crystals of the beetle toxin, the N terminus at the start of helix α_1 borders on a large solvent channel of about 30 Å diameter that crosses the unit cell along the *c* direction (Fig. 2*d*). This channel could allow access of sporulation-associated proteases to the cleavage site in parasporal crystals[19].

Receptor binding. The insecticidal selectivity of δ-endotoxins is due to high-affinity binding to specific membrane receptors[7-9,38], which in three cases seem to be glycoproteins[38-40]. For several δ-endotoxins the specificity-determining regions have been delimited by exchanging sequence segments between closely related toxins of differing specificities[13-15]. Guided by the location of secondary structures in the beetle toxin, a plausible alignment of δ-endotoxin sequences was made for the non-

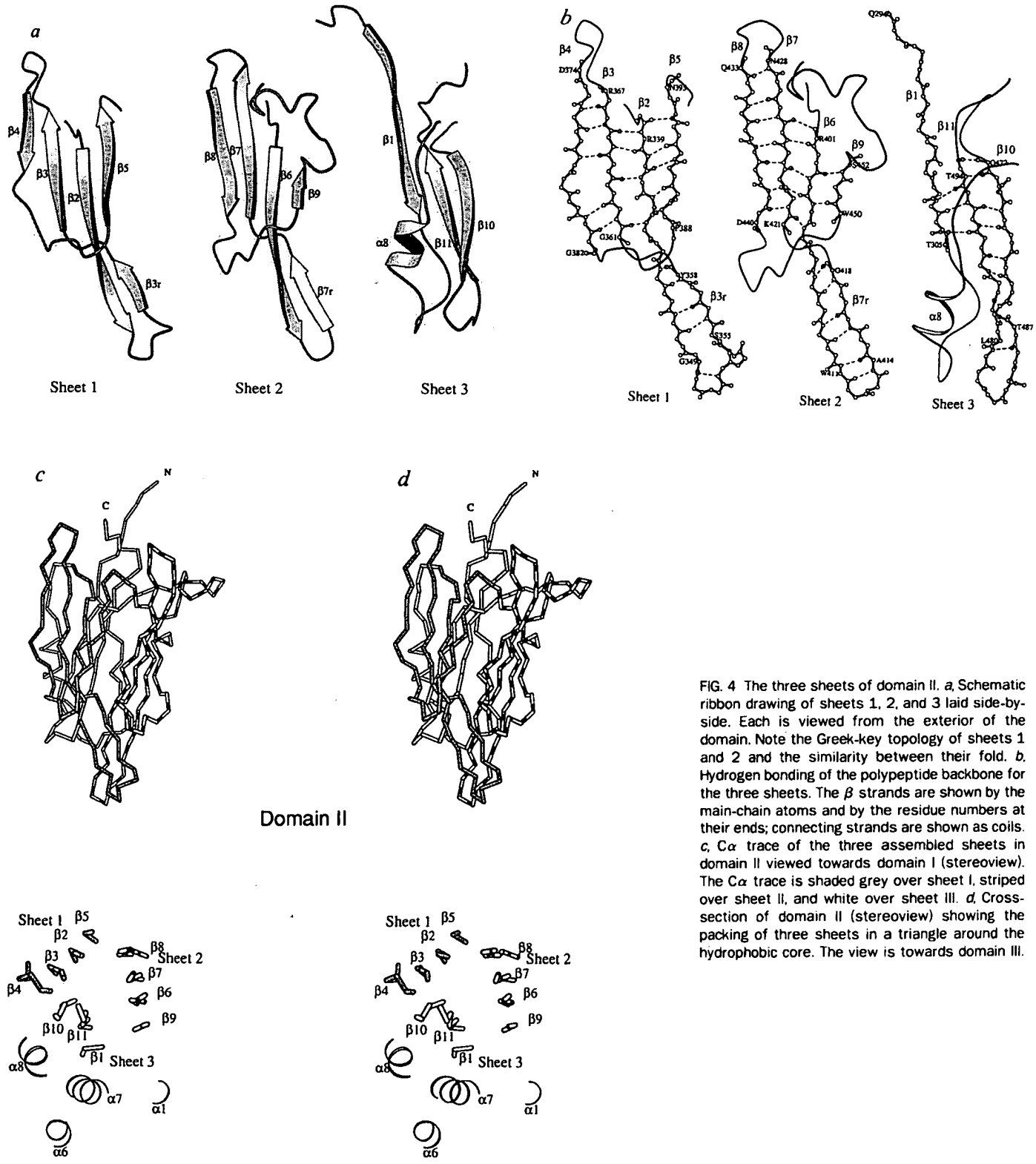


FIG. 4 The three sheets of domain II. *a*, Schematic ribbon drawing of sheets 1, 2, and 3 laid side-by-side. Each is viewed from the exterior of the domain. Note the Greek-key topology of sheets 1 and 2 and the similarity between their fold. *b*, Hydrogen bonding of the polypeptide backbone for the three sheets. The β strands are shown by the main-chain atoms and by the residue numbers at their ends; connecting strands are shown as coils. *c*, Cα trace of the three assembled sheets in domain II viewed towards domain I (stereoview). The Cα trace is shaded grey over sheet I, striped over sheet II, and white over sheet III. *d*, Cross-section of domain II (stereoview) showing the packing of three sheets in a triangle around the hydrophobic core. The view is towards domain III.

conserved regions (ref. 12, and T. C. Hodgman, unpublished results). Hence the genetically identified specificity-determining regions can be mapped to equivalent positions in the beetle toxin structure, and these fall mainly in domain II. For instance, the dual specificity of CryIIA for Lepidoptera and Diptera, as distinct from the Lepidoptera specificity in the closely related CryIIB, is determined by residues 307–382 of their sequences[14], which corresponds roughly to sheet 1 (Fig. 4*a*) plus strand β_6 in sheet 2 and the loop leading up to β_7, whereas the Lepidoptera specificity of CryIIB is dependent on a longer segment[14] that would include both inner strands of sheet 2. Similarly, the toxicities of CryIA(a) and CryIA(c) to two lepidopteran insects depend on three segments termed x, y and z (ref. 15): amino-acid substitutions in y can reduce toxicity by up to 2,000-fold, and segments x and y interact in determining specificity. Aligned with the beetle toxin structure, segment x corresponds roughly to the outer strands β_4 and β_5 of sheet 1 and the whole of sheet 2, including the loop entering β_{10} in sheet 3; y corresponds to

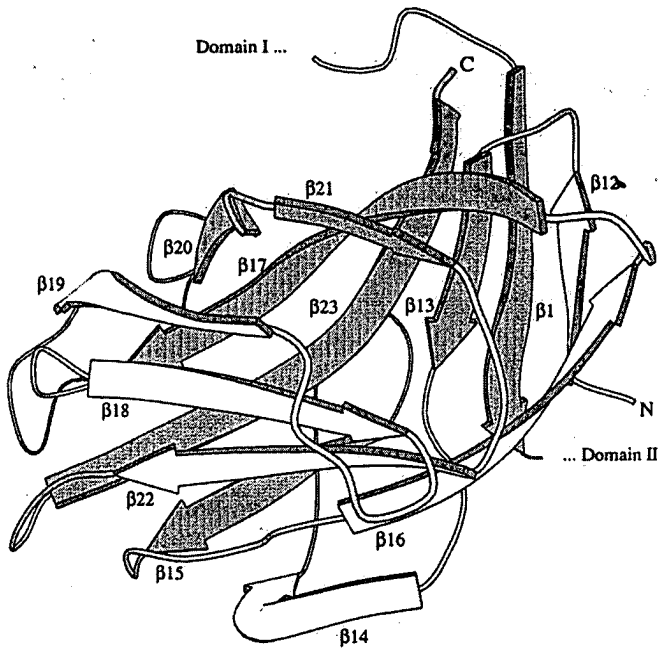


FIG. 5 Domain III, schematic ribbon representation of the β sandwich. β strands forming the inner sheet are shaded grey. The topology of an eight-stranded 'jelly-roll' can be seen by following the β hairpin starting with β_{13}, β_{15} and β_{23} in the inner sheet, continuing to β_{16} and β_{22} in the outer sheet, then β_{17} and β_{21}, β_{20} in the inner sheet, and ending with β_{18} and β_{19} in the outer sheet. β_{14} is an excursion from the hairpin and forms a fifth antiparallel strand of the outer sheet. Small parallel β sheets are added to one edge of the β sandwich, by hydrogen bonding of β_1 to β_{13} in the inner sheet and β_{12} to β_{16} in the outer sheet. Residue numbers in the β strands are: β_{12}, 502-506; β_{13}, 509-513; β_{14}, 519-525; β_{15}, 536-541; β_{16}, 547-554; β_{17}, 558-569; β_{18}, 573-579; β_{19}, 585-591; β_{20}, 604-609; β_{21}, 611-614; β_{22}, 619-625; and β_{23}, 631-643.

strand β_{10} of sheet 3 and the loop connecting β_{10} and β_{11}; and z extends from β_{11} to the C-terminal activation site. Furthermore, the interaction between x and y can be understood in terms of the proximity between β_4 on the edge of sheet 1 and β_{10} on the edge of sheet 3. Although z was inferred[15] to extend into domain III, the combined evidence from genetics and receptor-binding assays *in vitro* for Lepidoptera toxins[9,41] correlates receptor recognition with sequence variations within domain II. We note that the β ribbons from all three sheets terminate in loops in a small region on the molecular apex, in a manner reminiscent of the complementarity-determining region of immunoglobins.

Pore formation. The common mechanism of epithelial cell disruption by δ-endotoxins of widely different specificities is believed to be the formation of lytic pores of 10 to 20 Å diameter in the insect membrane[10]. The structure of the beetle toxin displays an apparatus for pore formation in the long, hydrophobic and amphipathic helices of domain I which could penetrate the membrane. Between the crystal structure in which the bouquet-like helical bundle internalizes all the hydrophobic surfaces, and the unknown pore structure where hydrophobic surfaces would be in intimate contact with the membrane lipids, large conformation changes must occur. In the absence of a full characterization of the pore-forming process, we propose the following by extrapolation from the crystal structure.

The trigger for the conformational changes may be provided by receptor binding and the consequent interaction of toxin with the membrane bilayer. Membrane insertion follows rapidly, so that a major part of the bound δ-endotoxin cannot be displaced from the brush-border vesicles by other toxins recognizing the same receptor sites[7,9]. As domain II and probably its apical region are most likely to bind the membrane receptors, the helices are expected to insert with the 'domain II end' (see Fig. 2*a*) oriented towards the cytoplasm. If helical hairpins are to initiate the membrane penetration, as probably happens for colicin[28,42,43], they will probably be linked at the domain II end. So either of the helix pairs α_6-α_7 or α_4-α_5 could be the likely initiator. The α_6-α_7 pair is favoured because it forms part of the conserved interface with domain II and is well positioned to sense the receptor binding. On the other hand, helix α_5 is the most conserved throughout the family of δ-endotoxins. Point mutations in α_5 reduce toxicity of a Lepidoptera toxin without reducing binding to membranes[44]. Proteolysis in the interhelical loops at the domain III end, as in the α_3-α_4 loop[19,32], may facilitate release of the helix pairs from the tertiary structure of the bundle. The insertion of a hairpin can create a defect in the membrane, allowing the rest of domain I to participate in pore formation in a cooperative manner. □

Received 22 July; accepted 19 September 1991.

1. Höfte, H. & Whiteley, H. R., *Microbiol. Rev.* **53,** 242-255 (1989).
2. Ellar, D. J. *et al.* in *Molecular Biology of Microbial Differentiation* (eds Hoch, J. A. & Setlow, P.) 230-240 (Am. Soc. Microbiol., Washington, DC, 1985).
3. Wilcox, E. R. *et al.* in *Protein Engineering: Applications in Science, Medicine and Industry* (eds Inouye, M. & Sarma, R.) 395-413 (Academic, New York, 1986).
4. Vaeck, M. *et al. Nature* **328,** 33-37 (1987).
5. Perlak, F. J., Fuchs, R. L., Dean, D. A., McPherson, D. L. & Fischhoff, D. A. *Proc. natn. Acad. Sci. U.S.A.* **88,** 3324-3328 (1991).
6. Barton, K. A., Whiteley, H. R. & Yang, N. S. *Plant Physiol.* **85,** 1103-1109 (1987).
7. Hofmann, C., Lüthy, P., Hütter, R. & Pliska, V. *Eur. J. Biochem.* **173,** 85-91 (1988).
8. Hofmann, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **85,** 7844-7848 (1988).
9. Van Rie, J., Jansens, S., Höfte, H., Degheele, D. & Van Mellaert, H. *Appl. Envir. Microbiol.* **56,** 1378-1385 (1990).
10. Knowles, B. H. & Ellar, D. J. *Biochim. biophys. Acta* **924,** 509-518 (1987).
11. Endo, Y. & Nishitsutsuji-Uwo, J. *J. Invertebr. Path.* **36,** 90-103 (1980).
12. Hodgman, T. C. & Ellar, D. J. *J. DNA Sequ. Map.* **1,** 97-106 (1990).
13. Ge, A. Z., Shivarova, N. I. & Dean, D. H. *Proc. natn. Acad. Sci. U.S.A.* **86,** 4037-4041 (1989).
14. Widner, W. R. & Whiteley, H. R. *J. Bact.* **172,** 2826-2832 (1990).
15. Schnepf, H. E., Tomczak, K., Ortega, J. P. & Whiteley, H. R. *J. biol. Chem.* **265,** 20923-20939 (1990).
16. Krieg, A., Huger, A. M., Langenbruch, G. A. & Schnetter, W. *J. appl. Entomol.* **96,** 500-508 (1983).
17. Höfte, H., Seurinck, J., Van Houtven, A. & Vaeck, M. *Nucleic Acids Res.* **15,** 7183 (1987).
18. Li, J., Henderson, R., Carroll, J. & Ellar, D. *J. molec. Biol.* **199,** 543-545 (1988).
19. Carroll, J., Li, J. & Ellar, D. J. *Biochem. J.* **261,** 99-105 (1989).
20. Arndt, U. W. *Meth. Enzym.* **114,** 472-485 (1985).
21. Wang, B. C. *Meth. Enzym.* **115,** 90-112 (1985).
22. Leslie, A. G. W. *Acta crystallogr.* **A43,** 134-136 (1987).
23. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, N. *Acta crystallogr.* **A47,** 110-119 (1991).
24. Brünger, A. T. *J. molec. Biol.* **203,** 803-816 (1988).
25. Chothia, C. & Finkelstein, A. V. *A. Rev. Biochem.* **59,** 1007-1039 (1990).
26. Chothia, C. *J. molec. Biol.* **105,** 1-15 (1976).
27. Burley, S. K. & Petsko, G. A. *Science* **229,** 23-28 (1985).
28. Parker, M. W., Pattus, F., Tucker, A. D. & Tsernoglou, D. *Nature* **337,** 93-96 (1989).
29. Richardson, J. S. *Adv. Prot. Chem.* **34,** 167-339 (1981).
30. Höfte, H. *et al. Eur. J. Biochem.* **161,** 273-280 (1986).
31. Choma, C. T. & Kaplan, H. *Biochemistry* **29,** 10971-10977 (1990).
32. Nicholls, C. N., Ahmad, W. & Ellar, D. J. *J. Bact.* **171,** 5141-5147 (1989).
33. MacIntosh, S. C., McPherson, S. L., Perlak, F. J., Marrone, P. G. & Fuchs, R. L. *Biochem. biophys. Res. Commun.* **170,** 665-672 (1990).
34. Schnepf, H. E. & Whiteley, H. R. *J. biol. Chem.* **260,** 6273-6280 (1985).
35. Adang, M. J. *et al. Gene* **36,** 289-300 (1985).
36. Sekar, R., Thompson, D. V., Maroney, M. J., Bookland, R. G. & Adang, M. J. *Proc. natn. Acad. Sci. U.S.A.* **84,** 7036-7040 (1987).
37. Chungjatupornchai, W., Höfte, H., Seurinck, J., Angsuthanasombat, C. & Vaeck, M. *Eur. J. Biochem.* **173,** 9-16 (1988).
38. Knowles, B. H., Thomas, W. E. & Ellar, D. J. *FEBS Lett.* **168,** 197-202 (1984).
39. Knowles, B. H. & Ellar, D. J. *J. Cell Sci.* **83,** 89-101 (1986).
40. Haider, M. Z. & Ellar, D. J. *Biochem. J.* **248,** 197-201 (1987).
41. Visser, B., Munsterman, E., Stoker, A. & Dirske, W. G. *J. Bact.* **172,** 6783-6788 (1990).
42. Lakey, J. H., Baty, D. & Pattus, F. *J. molec. Biol.* **218,** 639-653 (1991).
43. Song, H. Y., Cohen, F. S. & Cramer, W. A. *J. Bact.* **173,** 2927-2934 (1991).
44. Ahmad, W. & Ellar, D. J. *FEMS Microbiol. Lett.* **68,** 97-104 (1990).
45. Messerschimidt, A. & Pflugrath, J. W. *J. appl. Crystallogr.* **20,** 306-315 (1987).
46. Read, R. J. *Acta crystallogr.* **A42,** 140-149 (1986).
47. Kraulis, P. *J. appl. Crystallogr.* (in the press).

ACKNOWLEDGEMENTS. We thank P. R. Evans, A. G. W. Leslie and R. Henderson for advice and encouragement; K. Wilson and Z. Dauter for help with the image plate system; K. Nagai and P. J. McLaughlin for help in collecting film data; T. A. Jones for advice on model building; SERC Daresbury Laboratory and EMBL Outstation at DESY for use of synchrotron facilities; and T. Woollard and K. Hopkins for maintaining the rotating anodes. D.J.E. and J.C. acknowledge the support of the AFRC.